DOCUMENT RESUME

ED 253 428 SE 045 397

AUTHOR	Blosser, Patricia E., Ed.; Helgeson, Stanley L., Ed.
TITLE	Investigations in Science Education. Volume 10, Number 3.
INSTITUTION	ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, Ohio.; Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE	84
NOTE	77p.
AVAILABLE FROM	Information Reference Center (SMEAC/IRC), The Ohio State Univ., 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 (subscription $8.00, $2.75 single copy).
PUB TYPE	Collected Works - Serials (022) -- Reports - Research/Technical (143) -- Information Analyses - ERIC Information Analysis Products (071)
JOURNAL CIT	Investigations in Science Education; v10 n3 1984
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Academic Achievement; *Cognitive Development; *Cognitive Processes; *Concept Formation; Developmental Stages; Elementary Secondary Education; Higher Education; *Learning; Locus of Control; Problem Solving; *Science Education; *Science Instruction
IDENTIFIERS	*Science Education Research; Time Series Analysis

ABSTRACT

Presented are abstracts and abstractors' analyses of 10 studies related, in some manner, to learning. These studies focus on: (1) problem-solving processes used by experts and novices as they worked five mechanics problems (J. H. Larkin and F. Reif); (2) formal reasoning patterns of Italian high school students (M. L. Aiello-Nicosia, et al.); (3) the extent to which instructional and aptitude variable correlate with cognitive level (G. J. Pallrand and V. Moretti); (4) the development of correlational reasoning (A. Lawson et al.); (5) science learning when preadolescent children and their parents learn science content together (E. D. Gennaro et al.); (6) the relationship between formal reasoning ability and locus of control, academic engagement, and integrated process skill development (K. G. Tobin and W. Capie); (7) the development of hierarchical classification ability and the age at which this ability appears to be fully developed (W. E. Lowell); (8) relations of student, teacher, and learning environment variables to science learning attitudes (T. Haladyna et al.); (9) the effects to two small-group instructional strategies on children's nonvisual seriation abilities (M. J. Padilla and L. Ollila); and (11) the use of a time series design to examine concept development (V. Mayer and H. J. Kozlow). (JN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

SCIENCE EDUCATION INFORMATION REPORT

INVESTIGATIONS IN
SCIENCE EDUCATION

Volume 10, Number 3, 1984

THE ERIC SCIENCE, MATHEMATICS AND ENVIRONMENTAL EDUCATION CLEARINGHOUSE
in cooperation with
Center for Science and Mathematics Education
The Ohio State University

INVESTIGATIONS IN SCIENCE EDUCATION

Editor

Patricia E. Blosser
The Ohio State University

Associate Editor

Stanley L. Helgeson
The Ohio State University

Advisory Board

Anton E. Lawson (1984)
Arizona State University

John R. Staver (1986)
University of Illinois
at Chicago Circle

Gerald Neufeld (1984)
Brandon University

Glen S. Aikenhead (1986)
University of Saskatchewan

Richard Duschl (1985)
University of Houston
at University Park

National Association for Research in Science Teaching

Clearinghouse for Science, Mathematics, and
Environmental Education

INVESTIGATIONS IN
SCIENCE EDUCATION

Volume 10, Number 3, 1984

Published Quarterly by

The Center for Science and Mathematics Education
College of Education
The Ohio State University
1945 North High Street
Columbus, OH 43210

Subscription Price: $8.00 per year. Single Copy Price: $2.75.
Add $1.00 for Canadian mailings and $3.00 for foreign mailings.

Volume 10, Number 3, 1984

NOTES FROM THE EDITOR iii

Larkin, Jill H. and F. Reif. "Understanding and Teaching Problem-Solving in Physics." European Journal of Science Education, 1 (2): 191-203, April-June, 1973.
Abstracted by ELIZABETH KEAN 1

Aiello-Nicosia, M. L.; DiGiacomo; A. Machi; R. M. Spearandeo-Minoe; M. A. Allenza; and M. Zimmardi. "Study of Formal Reasoning Patterns by Means of Three Piagetian Tasks: Implications for Science Education." European Journal for Science Education, 2 (1): 67-76, 1980.
Abstracted by ANTON E. LAWSON 8

Pallrand, George J. and Victor Moretti. "Relationship of Cognitive Level to Instructional Patterns of High School Seniors." Journal of Research in Science Teaching, 17 (3): 185-190, 1980.
Abstracted by ANTON E. LAWSON 15

Lawson, A.; H. Adi; and R. Karplus. "Development of Correlational Reasoning in Secondary Schools: Do Biology Courses Make a Difference?" The American Biology Teacher, 41 (7): 420-425, 1979.
Abstracted by DAVID L. DUNLOP 21

Gennaro, Eugene D., Ann Bullock and Ada Alden. "Science Learning Experience Involving Adults and Their Preadolescent Children." Science Education, 64 (3): 389-296, 1980.
Abstracted by CONSTANCE M. PERRY 29

Tobin, K. G. and W. Capie. "Relationship Between Formal Reasoning Ability, Locus of Control, Academic Engagement and Integrated Process Skill Achievement." Journal of Research in Science Teaching, 19 (2): 113-121, 1982.
Abstracted by WILLIS J. HORAK 35

Lowell, Walter E. "The Development of Hierarchical Classification Skills in Science." Journal of Research in Science Teaching, 17 (5): 425-433, 1980.
Abstracted by DAVID F. TREAGUST And BARRY J. FRASER 41

Haladyna, T., R. Olsen, and J. Shaughnessy. "Relations of Student, Teacher, and Learning Environment Variables to Attitudes Toward Science." Science Education, 66 (5): 671-687, 1982.
Abstracted by MARCIA LAKES MATYAS 46

Padilla, Michael J. and Lloyd Ollila. "Effects of Small-Group Teaching on Acquisition and Transfer of Nonvisual Seriation Abilities." Science Education, 64 (3): 357-366, 1980.
Abstracted by GERALD G. NEUFELD 56

Mayer, Victor and M. James Kozlow. "An Evaluation of a Time-Series Single-Subject Design Used in an Intensive Study of Concept Understanding." Journal of Research in Science Teaching, 17 (5): 455-461, 1980.
Abstracted by FRANCES LAWRENZ 64

NOTES FROM THE EDITOR:

The ten articles reviewed in this issue relate, in some manner, to learning. Larkin and Reif studied problem solving processes used by experts and novices as they worked five mechanics problems. Aiello-Nicosia and colleagues conducted a descriptive study of the formal reasoning patterns of Italian high school students. Pallrand and Moretti investigated the extent to which instructional variables and aptitude variables correlate with cognitive level. Lawson et al. gathered descriptive data about the development of correlational reasoning. Gennaro et al. investigated science learning when preadolescent children and their parents learned science content together. Tobin and Capie studied the relationship between formal reasoning ability and locus of control, academic engagement and integrated process skill development. Lowell explored the development of hierarchical classification ability and the age at which this appears to be fully developed in the child. Haladyna and Shaughnessy examined learning environment variables as well as relations of students and teachers as these influenced attitudes toward science subject matter. Padilla and Ollila looked at the effects of two small-group instructional strategies on children's nonvisual seriation abilities. Finally, Mayer and Kozlow reported on the use of a time series design to examine concept development.

Patricia E. Blosser
Editor

Stanley L. Helgeson
Associate Editor

Larkin, Jill H. and F. Reif. "Understanding and Teaching Problem-Solving in Physics." European Journal of Science Education, 1 (2): 191-203, April-June, 1973.

Descriptors--*College Science, Engineering Education, Higher Education, Instruction, *Mechanics-Physics, Models, *Physics, *Problem-Solving, *Science Education, Skill Development, Teaching Methods

Expanded abstract and anlysis prepared especially for I.S.E. by Elizabeth Kean, University of Wisconsin-Madison.

Purpose

The stated purpose of this paper was to describe the authors' general approach to understanding problem solving processes in a science and how such understanding could be used to design instruction to improve student problem solving abilities. To illustrate their approach, they described the procedure used 1) to formulate models of problem solving processes used by an expert and by a novice to solve problems in physics, and 2) to test whether instruction based on the expert's strategies could be taught to novices to improve their problem-solving abilities.

Rationale

The authors believe that there is a need to formulate testable models of problem solving in use by novices and by experts. They assume that the novice and the expert approach problems differently. To gather information on these strategies, it is necessary to observe closely the processes of both the expert and the novice. From this information, they propose to formulate theoretical models of problem solving. Such models could be refined through repeated cycles of observation of actual problem solving behavior and analysis of the ability of the model to predict problem solving steps of each. Once the differences between expert and novice strategies are identified, they believe it possible to design problem solving instruction for novices, based on the expert's strategies.

The research described is in two parts: 1) Understanding the problem solving processes used by experts and novices, and 2) Designing instruction in problem solving.

Research Design and Procedure: Part I

Part I: The authors used a case study approach to identify strategies used by two subjects who each solved five mechanics problems. The expert was a physics professor who had recently taught mechanics; the novice was a successful student who had recently completed his first mechanics course. The novice was as successful as the expert in solving the five problems.

The raw data consisted of transcripts of tape recorded verbal comments of each problem solver as they worked through the five problems; these were combined with the written work of each. These verbal comments were the result of "thinking out loud." This record of the problem solving approaches of each were analyzed to produce "programs" (akin to computer programs) which produce, as their "output," steps in the problem solution. The programs were written as "production systems," specifying that whenever a particular condition is satisfied, a corresponding action is implemented.

The programs were refined by an iterative process. The predicted orders of steps were checked against the actual problem solving steps of each of the subjects and the programs were modified until a satisfactory match was obtained. The global features of the novice and expert problem solving approaches were given as Table 2.

Findings: Part I

The qualitative description of the "expert" and the "novice" problem solving strategies constituted the findings of Part I. Similarities were noted (e.g., both began by constructing an original description, with labeled sketch). Two major differences were found: 1) after the original description, the expert developed a low detail, qualitative "physical" description before beginning to construct a mathematical description, whereas the novice began immediately to construct a mathematical descrip-

tion; 2) the expert employed coherent methods, relating several principles into one functional unit, whereas the novice applied individual principles sequentially until he had succeeded in eliminating all unknown quantities.

Empirical support for these models was obtained from their ability to predict the sequences of steps used by each problem solver in solving all five problems, and by the clustering of written equations in time. The time interval for production of two sequential equations was random for the novice but clustered for the expert, indicating a rapid burst of equations within a relatively short period of time.

Research Design and Procedure: Part 2

After having identified two differences in the problem solving approach of the expert and the novice, two instructional sequences were designed, one of which incorporated "expert" strategies, the other "novice" strategies. Ten students in an introductory level physics course volunteered as subjects and were randomly assigned to learn "expert" or "novice" strategies. Each group was taught identical content (seven principles involving electric circuits) until they could apply each of the seven principles individually. Then each group was trained (equal time for each group) in either "expert" or "novice" strategies until they could verbally summarize the suggestions for problem solving which had been taught. The "expert" strategies included production of a low detail qualitative description and use of coherent methods, while the novice method stressed immediate use of mathematical descriptions and organization of equations according to quantities they contained (emphasis on sequential substitution into equation to solve for unknowns).

All students were then asked to solve three problems, each of which required application of several principles. Students worked individually, thinking out loud, with the experimenter pointing out errors but not suggesting solutions. Students were given a time limit on solving these problems.

Findings: Part 2

The five students who had been taught the expert strategies solved more problems (three students solved all three problems; the remaining two solved two) than did the students using novice strategies (one student solved three problems, the other four solved only one). The authors gave no evidence that either group of problem solvers actually used the types of approaches that they had been taught.

Interpretations

The first implications that the authors stated related to the teaching practices used in introductory science and engineering courses. Their research implied that it is important to teach students the two "expert" strategies of coherent methods and qualitative solutions. These were contrasted to the usual teaching methods in existing college courses which emphasize successive application of individual principles and excessive emphasis on premature precise mathematical formulations of problems. Instruction designed to enhance problem solving skills should encourge students' effective use of their natural verbal and pictorial arguments and how to translate them at the appropriate time into mathematical arguments.

Secondly, this work demonstrates the usefulness of this type of approach for studying the performance and teaching of complex intellectual skills. It is seen as an alternative to philosophical speculation or gross statistical investigations in classroom research. Such description of normative thinking approaches and development of testable teaching strategies based on those approaches can be combined with theoretical models to lead to advances in the science of education.

ABSTRACTOR'S ANALYSIS

This paper describes a method of obtaining information about the complex skill of problem solving. Its methodological roots are to be

found in Bloom and Broder (1950). Reif's group at Berkeley is in the forefront of research on understanding problem solving processes in science disciplines and using that understanding to develop more powerful instructional methods. A relatively recent summary of work in this field is found in Tuma and Reif (1980).

The first part of this paper described the general approach for exposing problem solvers' thought processes. By having them think aloud as they work on problems, the researcher obtains a record of the sequence of thoughts which arise along the path to a problem solution -- at least those thoughts of which the problem solvers are aware. The article includes only five examples of how the statements and written work of the problem solvers were translated into a generalized step in their solution processes (e.g., "locating forces"). No description was given of how these individual steps in a particular problem solution were collapsed into the generalized problem solving algorithms of the expert and novice. It is unclear what criteria were used to decide when the models were refined enough. What, exactly, constituted a match between a step in the model and the statements or written work of the problem solvers? Were there any statements which were impossible to categorize? Other researchers who would use this technique would not find the directions adequate for conducting their own studies.

A major part of this work involved the production of the two models of problem solving. The actual processes of problem solving are hidden inside the heads of the problem solvers. Researchers must be clever if they are to expose these hidden processes. The use of indirect supportive evidence such as the time clustering of equations used in solutions is a good example of how to build a case to support arguments when direct evidence is unobtainable. The second type of empirical evidence, agreement between the steps predicted by the models and the actual steps taken, could be strengthened by involvement of researchers other than the authors. Would colleagues who were not directly involved in determining the models also conclude that the models adequately predicted the problem solving steps used by the expert and novice? The authors seem to be asking us to take their word that the models fit.

It is possible that alternative interpretations of the raw data could be obtained. It is probable that the derivation of the generalized solutions required a high level of inference on the part of the analyst. Other analysts might derive problem solving models which attend to different features of the observed problem solving strategies. But what matter! The authors make no claim that their interpretations constitute the only interpretation of the data. Work in this area is still exploratory and we are still struggling to formulate concepts and processes which can form the basis for useful problem solving models.

The models which the authors develop seem reasonable. Yet, I believe it is important to keep in mind their tentative nature. We have no indication as yet as to how idiosyncratic these two problem solvers might have been, how unique their approaches to the particular set of problems, how those approaches might have changed if the problems had been easier, harder, set in a different area of mechanics or in a different discipline altogether. This study is an early one in the field. Much more work will be needed before the generalizability of these models can be demonstrated.

It is one thing to demonstrate that novices and experts differ in their approach to solving academic problems. It is quite another to determine whether the expert's strategies are appropriate for the novice. For example, the expert has at his/her command a vast repertoire of possible problem solving schemes. It may be that the need for a low detail qualitative solution is related to the need to select an appropriate strategy from among the multitude of possibilities. The novice, in contrast, has many fewer strategies at his or her disposal. At what point does the novice accumulate a sufficiently large repertoire of problem approaches so that a qualitative search is needed to select the most fruitful approach? Does merely teaching the novice to mimic the approach of the expert actually enhance the ability to the novice to solve classroom problems in physics?

The authors included a short section in which they described the problem solving success of ten novice students who had been instructed in either novice or expert problem solving strategies. However, the information included in the paper is not sufficient to attribute differential problem

solving results to the instruction provided (the group that had been instructed in use of the expert type strategies solved more problems than did the other group). Few details were provided as to the exact instruction provided for each small group of novices. Also, it was not possible to determine whether or not they actually used these methods, whether the students were similar in problem solving ability before instruction (random assignment of ten students to two groups does not insure that the groups are identical), whether increased success in solving the problems was affected by the time limit placed on the problem solvers, etc. Thus, the results in this section are merely suggestive and their tentative nature must be stressed. It would have been helpful for the authors to describe how some of the necessary variables could be controlled in designing future studies on explicit teaching of problem solving skills to novices.

In sum, this study provides a rationale for using a "case study" approach for examining detailed performance in a complex skill. It is an important study not because it provides definitive answers to well defined questions, but because it raises important questions that can help guide development of a complex and important area of research.

REFERENCES

Bloom, B., and L. Broder. Problem Solving Processes of College Students. Chicago: University of Chicago Press, 1950.

Tuma, D.T., and F. Reif, eds. Problem Solving and Education: Issues in Teaching and Research. Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers, 1980.

Aiello-Nicosia, M.L.; DiGiacomo; A. Machi; R.M. Spearandeo-Minoe; M.A. Allenza; and M. Zimmardi, "Study of Formal Reasoning Patterns by Means of Three Piagetian Tasks: Implications for Science Education." European Journal for Science Education, 2 (1): 67-76, 1980.

Descriptors--*Cognitive Development, *Cognitive Measurement, Educational Research, *Intellectual Development, Intelligence Tests, Learning Theories, *Physics, Science Education, *Scientific Concepts

Expanded abstract and analysis prepared especially for I.S.E. by Anton E. Lawson, Arizona State University.

Purpose

The study was purely descriptive. No hypotheses were explicitly stated nor tested. Three Piagetian tasks were individually administered to a sample of 300 Italian high school students, aged 14 to 18 years, randomly selected from four high schools with different approaches to institutions: humanistic, scientific, technological and pedagogical. The following questions were investigated: How consistent is intertask performance? What portion of students display concrete and/or formal reasoning on the tasks? Do sex differences in performance occur? Does performance vary with socio-economic status, chronological age, and type of school.

Rationale

Certain aspects of Piaget's theory are considered most helpful in accounting for differences in students' abilities to understand science concepts, conduct investigations and solve problems. In Italy, most high school physics courses tacitly (and presumbly incorrectly) assume that students are fully competent in all operations of formal reasoning. This leads to students who misunderstand all of the topics in the course. Teachers who are unaware of this problem tend to limit their objectives to the knowledge level only and "cook-book" labs,

rather than emphasing the formulating and testing of hypotheses. Thus study of the interactions between intellectual development and school achievement is in order as part of the process of reorganizing the physics content of high-school courses.

Three formal tasks were used in the present study on the assumption that each task can be considered a detector of a student's operational reasoning for a narrow class of equivalent problems.

Research Design and Procedure

The volume displacement task, chemical mixtures task, and the flexible-rods task were individually administered by three interviewers. Tasks were scored as follows: volume displacement (2A, 2B, 3A), chemical mixtures and flexible rods (2A, 2B, 3A, 3B). Students were grouped by grade. The first group corresponded to the 9th grade, the second to the 11th, and the third to the 13th grade (the final grade in Italian high schools).

Findings

On the displacement volume task 40% of the 9th graders, 48% of the 11th graders, and 58% of the 13th graders performed at the formal level. The corresponding percentages on the chemical mixtures task were 17%, 30% and 39% and in the flexible-rods task were 17%, 33%, and 41%.

Kendall's coefficient of concordance, which reflects the degree of association among the ranking of the tasks was .54 for the 9th grade sample, .61 for the 11th grade sample, and .58 the 13th grade sample (all $ps < .001$).

The Kruskal-Wallis test was applied to determine whether the tasks were of equivalent levels of difficulty. The chemicals and rods tasks were found to be of equivalent difficulty but the volume displacement task was found to be easier than the other task.

The 13th graders performed significantly better than did the 9th graders on the volume task. For the chemicals and rods tasks signficant differences were found between the 9th and 11th and between the 9th and 13th graders but not between the 11th and 13th graders.

Significant ($p < .05$) sex differences (in favor of males) were found on all tasks at all levels with the exception of the rods task for the 13th graders.

Type of school and socio-economic level (as assessed by occupation and education of parents) were not significantly related to task performance.

Interpretations

The finding that the volume task was easier than the other two tasks was interpreted to mean that formal reasoning is not an "all-or-nothing" affair. Students may use concrete reasoning on others. The difference in levels of performance suggests that it would be best to use a variety of tasks to obtain an accurate indication of reasoning level.

Male superiority on the tasks was interpreted as due to differences in upbringing or perhaps to differences in task familiarity. In Italy, differences in life style and education are evident. The female thinking is more frequently rooted in the practical and tangible, and this does not encourage them to think of possibilities and hypothetical situations.

The following educational implications were drawn: The stimulation of formal reasoning development represents a worthwhile aim of science education. Teaching strategies should include confronting students with ideas and materials at their own level and then gradually dealing with problems that require formal reasoning and directing students' attention to their own reasoning through discussion and reflection on what has been done and its implications. Preservice teachers should be trained to deal with students at varying levels of development.

ABSTRACTOR'S ANALYSIS

The findings of the present study with Italian students offers nothing really different that what has been previously found in a number of studies conducted in the United States and Great Britian. The problem of a mismatch between intellectual level of the student and the intellectual demands of science courses appears fairly universal. I could not agree more with the thrust of the authors' remarks concerning the need to use science to provoke the development of formal thinking. Unfortunately this advice often fails on the deaf ears of traditional high school science teachers, some of whom are barely formal operational themselves.

Determination of the cause or causes of sex differences on the formal tasks is certainly problematic. Of course differences in upbringing may be very important, yet genetic differences cannot be ruled out at the present time. This issue opens a can of worms much like the one opened by researchers interested in racial differences in intelligence. It is not an issue that it likely to be resolved at any time soon.

Although I agree with the stated educational implications, it should be pointed out that the prescription of effective teaching strategies to promote the development of formal reasoning lies beyond the scope of the present study. However, it does not lie beyond the scope of a number of other studies that have directly investigated ways of provoking formal thought development. To date, over 30 such studies have been conducted. Let us consider just two of these as examples of effective teaching procedures.

Schneider and Renner (1980) compared two methods of teaching physical science concepts to ninth grade students over a one semester period. One method, labeled formal instruction, followed a traditional pattern of lecture, motion pictures, filmstrips, textbook readings, questions and problems, supervised study and demonstrations. The second method, labeled concrete instruction, followed the exploration-invention-discovery learning cycle where emphasis is placed on

relatively open explorations, student data collection and teacher guidance. Results showed the concrete instruction method was superior to the formal method in content achievement on both immediate and delayed posttests. It was also superior in promoting intellectual development as assessed by a battery of formal tasks. The superiority of the concrete instruction group persisted on the delayed posttests (three months later). Thus support for the hypothesis that the learning cycle approach to teaching, due presumably to its ability to intellectually stimulate and challenge students' though processes, can not only improve their understanding of science content but can effect general advances in intellectual development and academic aptitude as well.

Wollman and Chen (1982) compared effectiveness of two classroom approaches for teaching fifth grade students how to control variables. One approach, called social interaction, involved demonstrating an event, (b) asking for a justification of the response in terms of evidence, (c) suggesting an alternative explanation and asking for an evaluation of it, and (d) accepting an explanation only after all other reasonable alternatives had been ruled out. These demonstrations and discussions lasted about 20 minutes in each of eight 45 minutes sessions. The remainder of each session was devoted to small group work with kits of materials which posed problems requiring the control of variables.

The other approach, called physical interaction, was identical to the previous approach except the initial social interaction sessions were omitted. Thus the physical interaction group received more time to work on the kit problems. Immediate and delayed posstests showed the two social interaction classes to be significantly better than the physical interaction class on control of variables problems. Using a stringent classification criteria, 70% of the social interaction students performed at the formal operational level while only 18% of the physical interaction students did. Thus, the authors concluded that their form of social interaction can have a substantial influence on students' performance above that produced by the weaker form of

physical interaction training used by traditional Piagentians (e.g., Kuhn and Anglev, 1976; Inhelder, Sinclair and Bovet, 1974). They suggest that a regular classroom teacher, with daily opportunities for engaging students in this type of social interaction discourse, should in the long run be able to revise substantially the level of functioning of all the pupils in the class.

These two studies are typical and support the present study's argument that the development of formal thought is not only a worthwhile but an obtainable objective. Given these results and given the fact that many students fail to develop formal reasoning abilities, under current conditions, the very clear educational implication is that many courses in many disciplines, over an extensive period of time (e.g. grades 4-12) should more directly concern themselves with the development of formal reasoning. Of course many decisions must be made to arrive at approaches reasonable for specific grade levels and specific disciplines, yet the training studies are very encouraging as they suggest that such efforts would pay off handsomely in improved reasoning ability and improvements in general academic achievement. The Science Curriculum Improvement Study's K-6 life and physical science program is a good case in point. The program developers paid very careful attention to the development of reasoning. Consequently, when the program is taught appropriately, students demonstrate measureable gains in intellectual development which show up as better understanding of the processes of science and as performance gains in mathematics, reading and social science (e.g., Linn and Thier, 1975; Renner, Stafford, Coffia, Kellogg, and Weber, 1973).

REFERENCES

Inhelder, B., H. Sinclair, and M. Bovet. Learning and the Development of Cognition. Cambridge: Harvard University Press, 1974.

Kuhn, D. and J. Angelev. "An Experimental Study of the Development of Formal Operational Thought." Child Development 47, 697-706, 1976.

Linn, M.D. and H.D. Thier. "The Effect of Experimental Science on Development of Logical Thinking in Children." Journal of Research in Science Teaching, 12: 49-62, 1975.

Renner, J.W., D.G. Stafford, W.J. Coffia, D.H. Kellogg, and M.C. Weber. "An Evaluation of the Science Curriculum Improvement Study." School Science and Mathematics, 73: 291-318, 1973.

Schneider, L.S. and J.W. Renner. "Concrete and Formal Teaching." Journal of Research in Science Teaching, 17: 503-518, 1980.

Pallrand, George J. and Victor Moretti. "Relationship of Cognitive Level to Instructional Patterns of High School Seniors." Journal of Research in Science Teaching, 17 (3): 185-190, 1980.

Descriptors--*Cognitive Development; Educational Research; Science Education; Science Instruction; Secondary Education; *Secondary School Science; *Student Interests; *Teaching Methods

Expanded abstract and analysis prepared especially for I.S.E. by Anton E. Lawson, Arizona State University.

Purpose

The purpose of the study was to determine the extent to which instructional variables, such as the number of years enrolled in high school studies, science, mathematics, and business courses, and aptitude variables, such as SAT math, SAT verbal, and IQ scores, correlate with cognitive level. Subjects were 100 recent high school graduates (average age = 18 years).

Rationale

Previous research has shown that many high school and college students do not reason consistently at the formal operational level. This presents a problem in terms of student performance in advanced science and mathematics courses, as many topics in these courses appear to require the use of formal reasoning patterns. A related issue concerns the extent to which science and/or mathematics courses can affect the acquisition of formal reasoning patterns. The present study sought to determine whether or not enrollment in science and/or mathematics classes is correlated positively with higher levels of formal reasoning.

Research Design and Procedure

A cross-sectional, no-treatment, static-group comparison design was used. Information on the number of years of couses taken in eight subject matter areas (1. social studies, 2. science, 3. mathematics, 4. language arts, 5. foreign language, 6. business, 7. industrial and home arts, and 8. art and music) was presumably obtained from student records, as were scores on SAT verbal, SAT math, IQ, and class rank. Cognitive level was measured through individual administration of three Piagetian tasks: (1) Combinations of Colorless Chemicals, (2) Equilibrium on the Balance, and (3) Combinations of Colored Tokens. Subjects were ranked on a five point scale on each task, ranging from preoperational to late formal operational performance. In addition to the three task scores, an overall cognitive level score was obtained by averaging the three task scores. Interrater reliability correlations based upon ratings of 15 randomly selected subjects exceeded .9 ($p < .001$).

Findings

Piagetian level ranged from early concrete to late formal operational and varied considerably on the three tasks. Scores were highest on the Chemical Combinations task and lowest on the Balance task. Intertask correlations ranged from .19 to .37. Only 11% of the Ss scored at the fully formal level.

Correlations of overall cognitive level with the eight subject matter areas were from lowest to highest as follows: business (-.24), art and music (-.21), language arts (-.09), social studies (-.03), foreign language (.12), industrial and home arts (.22), mathematics (.39), and science (.41). This pattern of correlations was accounted for primarily by differences in performance on the Balance task.

Overall cognitive level correlated at .43 with SAT verbal, .60 with SAT math, -.42 with class rank, and .55 with IQ.

A stepwise multiple regression analysis was performed in which the Piagetian task scores (designated the dependent variable) were regressed against the instructional and aptitude scores (designated as the independent variables).

The regression of the independent variables with the Balance scores yielded an R^2 value of .58, a .14 R^2 value with the Chemical Combinations score, a .16 R^2 with the Colored Tokens score, and a .46 R^2 value with the estimate of overall cognitive level.

Interpretations

Students operating at higher cognitive levels exhibit greater aptitude and take more courses in science and mathematics. The question remains as to whether differences in cognitive level influence curriculum choices or whether taking more courses in science and mathematics induces growth in cognitive level.

The ability to reason proportionally does appear to be related to the instructional pattern followed by students during high school. Many students appear to avoid science and mathematics courses because they lack the cognitive structures necessary to succeed. Courses emphasizing and integrating formal reasoning are appropriate, since formal reasoning patterns are so underdeveloped and appear to be related to instructional choice.

ABSTRACTOR'S ANALYSIS

Although I do not disagree with the thrust of the authors' interpretations, in some instances I fail to see how these follow from the data. In my opinion the only conclusion that is warranted is that performance on one task, the Balance Beam, correlated moderately with years of science and mathematics taken and negatively with years of business and art and music. The issue of cause and effect is not

addressed by the research design, thus does not allow Pallrand and Moretti to conclude, for example, that: "Many appear to avoid science and mathematics because they lack the cognitive structures necessary to succeed" (p. 189).

Subsequent research by Lawson and Bealer (1984) which reported reasoning performance prior to and following classroom instruction across grades 6 through 12 does, however, support the hypothesis that level of cognitive development influences choice of courses taken as opposed to the notion that taking additional science and mathematics courses induces cognitive growth. This, of course, is not to say that science and mathematics courses, if properly taught, cannot induce cognitive growth. It simply means that these courses, as presently taught, do not do so.

The lack of substantial correlation among the three tasks, as well as the lack of correlation of the Chemial Combinations and Colored Tokens tasks with the instructional variables, presents a problem that I believe was not adequately addressed by the authors. Apparently, (1) either these two tasks are not valid measures of cognitive level, or (2) the very notion of a unified level of advanced reasoning is not valid. Certainly many aspects of Piaget's model of formal thought, most centrally the hypothesis that formal thought is essentially isomorphic with propositional calculus (Inhelder and Piaget, 1958, p. 305), have come under considerable attack (cf. Wason and Johnson-Laird, 1972). Yet advanced reasoning patterns most assuredly exist and both statistically and theoretically seem to be linked (cf. Lawson, 1982a).

Further, and somewhat contrary to the results of the present study, formal reasoning ability has been found to show substantial positive correlations with standardized achievement measures (Iowa Tests of Educational Development) in reading (.69), language arts (.60), social studies (.72), as well as in science (.69) and mathematics (.70) (Lawson, 1982b). It would appear that the social studies and language arts courses taken by the students in the Pallrand and Moretti sample do not place a premium on reasoning, while the Iowa Tests of Educational Development in those areas do.

Use of stepwise multiple regression by the authors, although mathematically permissable, left me uneasy. In order to perform the regression analysis, the authors assumed the three Piagetian tasks to be the dependent variables while classroom instruction (i.e., the instructional and aptitude variables) were assumed to be independent variables. Certainly their data provide no reason to suspect that cognitive level is dependent upon classroom instruction. And, as mentioned, the Lawson and Bealer (1984) study indicates that the relationship is probably the other way around. Thus, the regression analysis, in my view, served no useful purpose and only confused the issue.

Another problem arises in the data reporting, or rather in the lack of it. The authors failed to report the complete matrix of correlation coefficients. Therefore, it was not possible to determine, for example, whether SAT mathematics performance correlated more highly than the Balance task performance did with the number of math courses taken, or to what extent IQ correlated with class rank or with SAT mathematics and SAT verbal performance. These correlations would have been interesting in light of the fact that class rank was found to correlate negatively with the Balance task! Does this mean that good reasoning skills are a detriment to getting good grades? Does getting good grades require one to <u>not</u> question, memorize, and blindly do as one is told even if one has no idea why? This is a major issue and one that was not addressed by the authors.

Allow me to make one final point concerning adequacy of the written report. The authors used abbreviations for each of the variables investigated (e.g., Equilibrium on the Balance became "Bal," Combinations of Colored Tokens became "Per"). These abbreviations were used in the body of the paper as well as in the tables. This I found particularly awkward. The paper was relatively short and the tables contained ample space to completely spell out the names of the variables. This would have saved the reader the considerable time spent in flipping back through the paper to recall what the abbreviations stood for.

REFERENCES

Inhelder, B. and J. Piaget. The Growth of Logical Thinking from Childhood to Adolescence. New York: Basic Books, 1958.

Lawson, A. E. "The Nature of Advanced Reasoning and Science Instruction." Journal of Research in Science Teaching, 19 (9): 743-760, 1982a.

Lawson, A. E. "Formal Reasoning, Achievement, and Intelligence: An Issue of Importance." Science Education, 66 (1): 77-83, 1982b.

Lawson, A. E. and J. M. Bealer. "The Acquisition of Basic Quantitative Reasoning Skills During Adolescence: Learning or Development?" Journal of Research in Science Teaching, 21 (4): 417-423, 1984.

Wason, P. C. and P. N. Johnson-Laird. Psychology of Reasoning: Structure and Content. Cambridge: Harvard University Press, 1972.

Lawson, A.; H. Adi; and R. Karplus. "Development of Correlational Reasoning in Secondary Schools: Do Biology Courses Make a Difference?" The American Biology Teacher 41(7): 420-425, 1979.

Descriptors--*Biology; *Cognitive Development; Developmental Psychology; Developmental Stages; *Educational Research; *High School Students; Logical Thinking; *Mental Development; Secondary Education; Secondary School Science

Expanded abstract and analysis prepared especially for I.S.E. by David L. Dunlop, University of Pittsburgh at Johnstown.

Purpose

The purposes of this study were (1) to provide an initial body of descriptive data concerning the development of one component of scientific reasoning - correlational reasoning - during the secondary school years; and (2) to determine what, if any, effect taking a standard course in high school biology has on this development.

Rationale

The rationale for this study is based on portions of Piaget's developmental theory. More specifically, the authors indicate that, within Piaget's theory, correlational reasoning is considered to be a formal stage acquisition due to its hypothesized dependence on the development of formal operations (Inhelder and Piaget, 1958). Further, it is generally agreed that correlational reasoning is developmentally advanced as it involves understanding of relations between previous relations (Lunzer, 1965), acceptance of ambiguity as a starting point for further reasoning (Collis, 1972), and second-order operations (Lovell, 1971).

Research Design and Procedures

Five hundred and seven students enrolled in grades 6, 8, 10, 12, and college freshmen and sophomores were the subjects in this study.

Selection of the participants was not randomized; however, students were selected from schools within the same socio-economic level at each grade.

The experimental design was a modification of the Time-Series quasi-experimental design described by Campbell and Stanley (1966). The authors used standard notation to represent their design as follows:

$$O_1 \quad O_2 \quad X \quad O_3 \; O_4 \; O_5$$

Where O_1 represents the test scores (observations) of the sixth-grade subjects, O_2 represents scores of the eighth grade subjects, etc. The X represents the treatment (taking a two semester course in biology). The assumption was made that all of the twelfth grade and college students had taken a high school biology course.

Two pencil-paper puzzles of correlational reasoning (Adi, Karplus, Lawson, & Pulos, 1978) were used to test each student. The puzzles required approximately 10 minutes to complete and were part of a larger battery of items which utilized the entire class period. Five response categories were developed, two of which represented successful solutions and three of which represented unsuccessful solutions to the puzzles. The unsuccessul solutions were classified as non-correlation reasoning while successful solutions were considered to represent correlational reasoning. Each response was scored by at least two raters, and agreement among raters was approximately 90%. Differences were resolved by discussion.

Findings

The results of the data analysis indicate that correlational reasoning was used by 3-4% of the sixth grade students and its use increased to approaximately 60-65% in the college group. Most responses (68.3%) were classified as illogical. The percentage of illogical responses decreased gradually up to the college-age subjects; however, even 22.1% of this group gave responses which were considered illogical. The authors point out that no abrupt increase in the use

of correlational reasoning was found between the eighth and tenth grade samples. The data from these two puzzles indicate that the biology course had no positive effect on the students' abilities to use correlational reasoning.

The finding that correlational reasoning was used by only 3-4% of the sixth grade students and that its use increased gradually to about 60-65% of the college students indicates that correlational reasoning does, for a majority of the students, "develop" during adolescence as developmental theory states.

ABSTRACTOR'S ANALYSIS

Several studies have suggested a positive relationship between formal reasoning and achievement (Keating, 1976; Cloutier and Goldschmid, 1976); hence, the basic tenet of this study is of obvious importance. It should also be recognized that the research has frequently shown that a significant percentage of high school and college students do not operate at the formal level of operations (McKinnon and Renner, 1971; Tomlinson-Keasey, 1972; Chiapetta, 1974; Lawson, 1974). While it is probably true that much of the concrete operational thought is a result of the students' inability to operate at a higher level, it has been shown that in some instances a student's level of reasoning is often below his or her capacity. Furthermore, his/her preference rather than his/her ability determines what method will be selected to solve a problem (Dunlop and Fazio, 1976).

Another aspect to consider in this type of research is the theory which underlies the rationale of the project. In this case the underlying rationale is based upon Piaget's developmental theories. Although these theories are generally accepted by researchers in the field, Linn (1981) cites several areas of debate in this area of research. First, there have been some questions raised concerning formal reasoning as a construct different from general ability and achievement (Keating, 1976; Linn, 1978). Second, researchers have suggested that Piaget's theory may not accurately account for empirical findings (Ennis, 1975; Lawson, 1976; Levine & Linn, 1977; Siegler, 1976), and finally, nonuniversality

in attainment of formal thought, as measured by Piagetian tasks, has been demonstrated (Blasi and Hoeffel, 1974).

In a recent article by Lawson (1983) a few studies were summarized in which the issue of the role of propositional logic in formal reasoning was addressed. Two key results were reported: (1) many students who succeed on proportions, probability, and correlation items failed the propositional logic items, and (2) a principal-component analysis of the items found two distinct factors with the formal reasoning items loading on the first factor and the propositional logic items loading on the second factor. Lawson's conclusion was that propositional logic and formal reasoning were distinct intellectual abilities. Treagust (1979) and Lamb & Betkouski (1980) both disagree with Lawson's conclusions. Treagust considers propositional logic as a fundamental part of formal reasoning, while Lamb and Betkouski suggest that propositional logic is a more sophisticated stage of intellectual development and develops after formal operations.

The conclusion of this study, that the biology course had no positive effect on the correlational reasoning of the students, is both refuted and supported by other research. The conclusions in a study by Lawson and Snitgen (1982) appear to contradict the conclusion in this article. They examined the effect of a one-semester college biology course on the development of students' ability to reason formally. Their data indicate that an inquiry-based classroom that includes specific provisions for the instruction of a variety of components of formal reasoning does produce measurable advances in their use. However, the extent to which this finding was influenced by a test-retest effect and/or a ceiling effect was uncertain. Further, McKinnon and Renner (1971) reported that college freshmen made significant gains in the development of formal reasoning after taking a one-semester, inquiry-oriented science course; but when McKinnon and Renner's study was replicated by Blake and Nordland (1978), they did not find similar development by their students. It should be mentioned that Lawson's study was similar to both McKinnon & Renner and Blake & Nordland; however, Lawson included specific provisions for instruction in the control of variables, and proportional, correlational, probabilistic, and combinatorial reasoning. Lawson's study also differed from previous studies

in that he assessed the student's general intelligence and degree of field independence on trainability. In support of this study are other conclusions by Lawson (1983). He examined the role of the biconditional in the acquisition of formal operational schemata during adolescence and concluded that formal operations develop in an increasing number of students during adolescence, yet their development appears far from universal and not as a consequence of any specific curricular innovations at any particular grade.

The authors described their experimental design as a novel modification of the Time-Series quasi-experimental design described by Campbell and Stanley (1966). Although possibly a sound design in certain situations, it was weakened in this instance by the need to assume that the biology courses of the 10th grade students, 12th grade students, and the college students were equivalent. It seems that some assurance must be given to the reader that the treatment (biology course) was roughly equivalent for each group. Further weakening resulted from the lack of randomized samples. This is especially true for the college group because it is doubtful if this group represents the general population in a manner similar to the sixth and eighth grade groups.

Since the study focused on the effects of a biology course, it would have been useful to know whether or not the subjects received a junior high course in biology as is frequently the case in many school districts. In retrospect, this appears to be a moot point; however, a brief discussion of this point seems warranted in light of the relatively good success by the 8th grade group in comparison to the 10th grade group.

Finally, the conclusion that the authors reached (taking a standard two-semester high school course in biology had no measurable effect on the development of correlation reasoning) was based on a very small number of biology courses. In the case of the treatment group, it seemed that this sample size was just one. Is it possible that different results could have been found in other environments with different texts and/or different teachers?

In summary, it is the opinion of this reviewer that the questions raised in this article are indeed relevant to science education and that the authors did achieve their stated purposes; however, because of the

complex nature of this area of research and because of the conflicting conclusions in the literature, more work must be done before one can claim to fully understand the processes which lead to the development of correlational reasoning in all students.

REFERENCES

Adi, H. et al. "Intellectual Development Beyond Elementary School VI: Correlational Reasoning." School Science and Mathematics, 78(8); 675, 1978.

Blake, A. & F. Nordland. "Science Instruction and Cognitive Growth in College Students." Journal of Research in Science Teaching, 15(5); 413-419, 1978.

Blasi, A. & E. C. Hoeffel. "Adolescence and Formal Operations." Human Development, 17; 344-363, 1974.

Campbell, D. T. & J. C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1966.

Chiapetta, E. L. "A Perspective on Formal Thought Development." Paper presented at the National Association for Research in Science Teaching Convention. Los Angeles, 1975.

Cloutier, R. & M. Goldschmid. "Individual Differences in the Development of Formal Reasoning." Child Development, 47; 1097-1102, 1976.

Collis, K. F. A Study of Concrete and Formal Operation in School Mathematics. Unpublished doctoral dissertation. University of Newcastle, New South Wales, Australia, 1972.

Dunlop, D. L. & F. Fazio. "Piagetian Theory and Abstract Preferences of College Science Students." Journal of College Science Teaching, 5(5); 297-300, 1976.

Ennis, R. H. "Children's Ability to Handle Piaget's Propositional Logic: A Conceptual Critique." American Educational Research Journal, 45; 1-41, 1975.

Inhelder, B. & J. Piaget. Growth of Logical Thinking From Childhood to Adolescence. New York: Basic Books, 1958.

Keating, D. P. "A Piagetian Approach to Intellectual Precocity." In D. P. Keating (Ed.), Intellectual Talent: Research and Development. Baltimore: John Hopkins University Press, 1976.

Lawson, A. E. "Relationship of Concrete and Formal Operational Science Subject Matter and the Developmental Level of the Learner." Paper presented at the National Association for Research in Science Teaching Convention. Chicago, 1974.

Lawson, A. E. "Formal Operations and Field Independence in a Heterogeneous Sample." Perceptual and Motor Skills, 42; 981-982, 1976.

Lawson, A. E. "The Acquisition of Formal Operational Schemata During Adolescence: The Role of the Biconditional." Journal of Research in Science Teaching, 20(4); 347-356, 1983.

Lawson, A. E. & D. A. Snitgen. "Teaching Formal Reasoning in a College Biology Course For Pre-service Teachers." Journal of Research in Science Teaching, 19(3); 233-248, 1982.

Levine, D. I. & M. C. Linn. "Scientific Reasoning Ability in Adolescence: Theoretical Viewpoints and Educational Implications." Journal of Research in Science Teaching, 14(4); 371-384, 1977.

Linn, M. C. "Formal Operations: Construct or Conglomerate?" Formal Operator, 1(4); 2-4, 1978.

Linn, M. C., S. Pulos, & A. Gans. "Correlates of Formal Reasoning: Content and Problem Effects." Journal of Research in Science Teaching, 18(5); 435-447, 1981.

Lovell, K. "Some Problems Associated With Formal Thought and its Assessment." In D. R. Green et al (Eds.), Measurement and Piaget. New York: McGraw-Hill, 1971.

Lunzer, E. A. "Problems of Formal Reasoning in Test Situations." In P.H. Mussen (Ed.), European Research in Child Development Monographs, 30(2): 19, 1965.

McKinnon, J. W. & J. W. Renner. "Are Colleges Concerned With Intellectual Development?" American Journal of Physics, 39: 1047-1052, 1971.

Siegler, R. S. "Three Aspects of Cognitive Development." Cognitive Psychology, 8: 481-520, 1976.

Tomlinson-Keasey, C. "Formal Operations in Females from Eleven to Fifty-four Years of Age." Developmental Psychology, 6(2): 364, 1972.

Treagust, D. F. "Comments on 'The Acquisition of Propositional Logic and Formal Operational Schemata During the Secondary School Years'." Journal of Research in Science Teaching, 16(4): 363-366, 1979.

Gennaro, Eugene D., Ann Bullock and Ada Alden. "Science Learning Experience Involving Adults and Their Preadolescent Children." Science Education, 64 (3): 389-296, 1980.

Descriptors--Adolescents; Adults; *Animal Behavior; Community Programs; *Family Programs; *Nonschool Educational Programs; *Science Education; Science Programs; Secondary Education; *Zoos

Expanded abstract and analysis prepared especially for I.S.E. by Constance M. Perry, University of Maine at Orono.

Purpose

The study researched the following questions:

1. Is it possible for preadolescent and/or adolescent children and their parents to learn science content together so that it is a valuable experience for both children and their parents?
2. What are their attitudes towards learning in this manner?

Rationale

Although there are many studies on children learning science, adults learning science, and some on programs teaching adults parenting skills or helping parents to deal with certain academic skills of children, very few studies are reported on programs that focus on parents and their preadolescent and/or adolescent children learning science together in a format that encourages interaction of family members. It was believed that the structuring of time for parents and their children to do things together, in this case learn science, would be of value to the families. The study was designed to evaluate that premise.

Procedures and Research Design

Course. A "formal" (authors' term) course in animal behavior was offered for parents and their middle school aged children at an informal learning center, the Minnesota Zoological Garden (MZG). The course met for three hours on five consecutive Saturdays. The presentation of material was aimed toward the adolescent. An inquiry mode was used. From 15 to 30 minutes of each class were used in comparing and contrasting animal and human behaviors as it pertained to family structure and communication. In addition to the classes at MZG, each family was given a male and female gerbil for home study. Home activity packets were provided which contained instructions for observations and questions parents and children were to complete together.

Sample. Twenty-five volunteers attended the classes but the final sample consisted of 18; ten adults and eight children.

Design and instrumentation. A test consisting of 15 multiple choice items was administered as a pre and post test. The test attempted to measure cognitive understanding of animal behavior. The remainder of the instrumentation was administered at the end of the course only.

A questionnaire designed to assess how they felt about the shared learning experience was completed by all participants at the last class. A five-point scale (5, a great deal to 1, none) was used by the sample to describe how much they learned about each topic (environmental pressures, observational skills, aggressive behavior, imprinting, territorialty, and learning).

Another set of items asked participants to agree or disagree with statements that focused on the usefulness of class and home activities in increasing knowledge of animal behavior and enhancing communication in the home. In addition, participants responded to an open-ended request to comment on one particularly rewarding experience as a result of the course.

Data analysis. A t-test was used to compare pre and post test scores of children, adults, and total participants on the multiple-choice test. Percentages of responses were figured for each point of the 5-point scale on the questionnaire and a simple tally was kept as to the number who agreed or disagreed on the usefulness of class and home activities in increasing knowledge and family communication. The results of the open-ended request for a rewarding experience were categorized into three groups of responses: family-oriented, science content-oriented, and both.

Findings

Pre and post 15 item multiple-choice test. Differences in scores for the children, the adults, and all participants were significant at the 0.01 level of probability. Learning, as measured by the test, did take place.

Questionnaire on shared learning experience. Eighty-seven of the responses indicated that the people believed their knowledge had increased some (scale point 3) or better than some (scale point 4 or 5), and 26% of the responses were marked "a great deal" (5) in all six topic categories. The areas where people believed they had learned the most were observational skills, aggression, and territoriality, which were topics stressed in class and the home activity packets.

Items on the usefulness of class and home activities in increasing knowledge and communication. All agreed that the class and home activities increased their understanding of animal behavior. In addition, all agreed communication in their homes increased as a result of the home activities and all but one adult participant agreed that communication increased at home as a result of the classes.

Open-ended request for rewarding experience. All responses were included in the Appendix. Proportionately more adults than children gave a family-oriented response as being the particularly rewarding experience.

The results were:

Family-oriented responses

2 (adult)

Content-oriented responses

4 (adult)

6 (children)

Family-and-content-oriented responses

3 (adult)

3 (children)

Interpretations

Courses offered to parents and children should be considered supplementary to school offered courses and should probably be on topics in which both adolescents and parents have limited knowledge so that they can learn together.

The five week time period seemed a good length although parents said they wished the course were longer. The age group of the children seemed ideal because the middle school years may be important for raising science interest and career awareness in order that the students will elect science courses in high school.

ABSTRACTOR'S ANALYSIS

The Gennaro, Bullock, and Alden study is a beginning in a relatively new area of study. They chose to see if parents and children could learn science together and have that learning be "a valuable experience" for all family members involved. The project was termed a pilot study and as such it answered a few questions--that families can learn together, believe they can learn together, and also believe the course increased family communication. The study also raised several more questions for research, some of which will be

brought up later in this analysis.

It was a little difficult to follow the procedure for the study. The research questions were not set apart from the text and some of the procedures used were interspersed with the findings. It would have been easier to follow if all procedures were in one section, including instrumentation, and only the findings were stated in the Results section.

The sample was very small and consisted only of volunteers who expressed an interest to learn science as a family. Both the size and the selection are factors that could have important effects on the study's results, but were not mentioned by the authors.

The course topics, class activities, and the home activity packets were clearly described. The reader could easily visualize the type of class and what the participants were to do together at home.

The multiple choice content test, as stated by the authors, turned out to be too easy. A more difficult test would obviously remedy that weakness. It would have been worthwhile to have validity and reliability scores for the instrument. They could have added strength or caused one to question the finding that significant learning did occur. Other instrumentation seemed to measure (1) if participants believed they learned; (2) if they believed family communication increased; and (3) what type of course experience was the most rewarding. One must assume that these measures are what the authors meant in the research questions when they stated that the learning be a valuable experience and that attitudes toward learning as a family would be measured. From the authors' description of the instrumentation, other than the content test, it appears that only when they asked what experience was most rewarding and tallied the family-oriented responses were they measuring attitude toward learning together. The questions that asked participants how much they believed they learned showed that the participants believed they learned much, but that doesn't appear to assess directly how the participants felt about shared learning, which is the authors' stated purpose for those items. The questions on communication one must assume fit under the

research question stating that learning together will be a valuable experience. The study would be clarified if the term 'valuable experience' was defined to include increased communication and a belief that learning occured in the shared environment.

Under the Results section, the authors state that "Parents and their children were observed to share responsibilities when working together... At the first sessions, adults tended to respond equally so." These statements go beyond the research questions of this study and are more observation than result, but, if researched, each of the two statements could make a study in itself and add important findings in the area of family-shared learning.

Also, in the Results section, the authors state that proportionately more adults than children provided a family-oriented response as being the particularly rewarding experience. One must look in the Appendix to find out that only two responses were family-oriented, both made by adults, although six were combination responses (family and content) made by three adults and three children. Numbers in the text would have clarified the findings.

The Discussion section makes several suggestions for further research such as checking attitudes toward science pre and post, and studying the contributions of the family component to learning and attitudes of the participants. They also suggest exploring subject areas other than science. In addition, the authors speculate that such courses could predispose students to elect science courses in high school. Research in that area would be most worthwhile. Perhaps studying attitudes toward life-long learning and childrens' view of their parents as learners would be another area of interest for research.

Obviously much must be deleted from a study when it is reported in a journal article, however, a brief section at the end of the article summarizing the findings and drawing conclusions would have clarified the written report. Withstanding the criticisms made here, the authors have chosen a valuable area of study and by conducting the study have opened up a relatively new area of research for further exploration.

Tobin, K. G. and W. Capie. "Relationship Between Formal Reasoning Ability, Locus of Control, Academic Engagement and Integrated Process Skill Achievement." Journal of Research in Science Teaching, 19 (2): 113-121, 1982.

Descriptors--*Abstract Reasoning; Elementary School Science; Elementary Secondary Education; Grade 6; Grade 7; Grade 8; *Locus of Control; Measures Individuals; *Predictor Variables; *Process Education; *Secondary School Science; Science Education; *Student Behavior; Student Characteristics; Time on Task

Expanded abstract and analysis prepared especially for I.S.E. by Willis J. Horak, University of Arizona.

Purpose

This study analyzed learner attributes and behaviors that were conjectured to be related to science process skill achievement. The learner attributes of formal reasoning ability and locus of control were chosen since previous studies have indicated that these variables are important predictors of certain other types of student achievement. Previous studies have also indicated that academic engagement is related to student achievement. This study attempted to clarify the previously reported observed relationship by analyzing specific types of academic engagement.

Rationale

Science educators have for many years indicated that one objective of teaching science in the elementary and middle schools should be process skill acquisition. However, there is still a need to determine specific student characteristics and behaviors which might enhance process skill achievement. Prior studies have often led to results which are not directly generalizeable to this achievement area.

Logic would seem to indicate that formal reasoning ability, locus of control, and academic engagement should be related to process skill achievement. The children who are formal operational should be more able tc hypothesize, control variables, and define operationally than the children who are not formal operational. Similarly, children who are internal should be more able to engage in data gathering activities and make greater use of information than students who are external. Lastly, students who use integrated processes during science activities should have higher process skill achievement than those who do not use these skills.

Research Design and Procedures

The study sample consisted of students chosen from 13 sixth, seventh, and eighth grade intact classes. Based upon formal reasoning scores, a stratified random sample of twelve students was chosen from each classroom. The students were given the Intellectual Achievement Responsibility Scale to measure locus of control, the Test of Logical Thinking to measure formal reasoning ability, and a self developed group test of integrated science process achievement as both a summative and a retention criterion measure. The retention measure was given four weeks after the summative measure.

Academic engagement was rated using the Transactions in Science direct observation instrument. This instrument contains categories of behavior that are logically related to the acquisition of integrated science processes and also additional categories of behavior that frequently occur in science classrooms. Behavior was observed and coded every thirty seconds in a randomly determined four second interval. For analysis, these observations were pooled over all observations. The behavior coding took place during a two week period in which eight science lessons, whose objectives included process skill acquisition, were being taught.

Findings

Correlations of the dependent and independent variables revealed significant relationships between process skill achievement, both summative and retention, and attending behaviors, generating behaviors, and formal reasoning ability. A regression procedure used to determine the variance in science process achievement attributable to these measures obtained an R-Square value of 0.40 for the summative measure and of 0.48 for the retention measure. In both, formal reasoning accounted for most of the variance with a value of 0.32 for the summative measure and of 0.41 for the retention measure.

Interpretations

The findings once again show formal reasoning ability is an important predictor of science achievement. However, achievement increases might be effected by increasing pupil attending and generalizing behavior. Since locus of control was related to attending behavior, knowledge of locus of control might make it possible for teachers to maintain better involvement with external students.

The study also found that the proportion of time students spent in objectives-related behavior was often less than one percent. The study may have found higher correlations with the academic engagement categories and the criterion measure if there had been a higher frequency of occurrences of the described behaviors. Thus, logical relationships between these behaviors need to be studied further. Effects on achievement could be more accurately determined if a suitable range was obtained in each category.

ABSTRACTOR'S ANALYSIS

As indicated in the expressed rationale for this research, there is a need for coordinated studies in science education that specifically link student process skill achievement to both student characteristics and student behavior at the middle school level. There have been studies that attempt to relate process skill achievement and student characteristics at other levels and these give some insights into this area of research. One such study, for example, is that utilizing preservice elementary education majors (Campbell and Perez, 1977). Others are easily locatable in the published research. This study is a good start in that direction for middle school pupils. The study did uncover some very interesting relationships and some possible problems with this area of research. Before I discuss these, however, I would like to make a few comments on the overall reporting of the study.

First of all, I would have liked to see a table of means and standard deviations for each of the measures. Without this many questions come to mind which cannot be directly answered. For example, what was the relationship between the criterion measure scores on the summative and the retention test? Formal reasoning ability accounted for a higher percentage of the variance on the retention test. Was this possibly due to a lower score overall or could there have been some other interracting variables? What were the characteristics of the locus of control scores? Although it is stated that there is evidence for no systematic relationship between formal reasoning ability and locus on control, could it be possible that the sampling procedures did affect this variable? If the standard deviation of the scores is low, there may yet be an undetermined relationship. Similar questions may also be asked in relationship to the formal reasoning scores. Previous studies have shown that sixth, seventh, and eighth grade students do not score high on formal reasoning tests. In fact, in their own validation studies on the Test of Logical Thinking, the researchers found approximately 3% and 4% of sixth and seventh grade

students, respectively, score higher than 3 (Tobin and Capie, 1981). Did the researchers have to pick all the students from these groups who did score high? Why did they pick the reference points of two and three on this specific test?

Similarly, a table of the actual percentages of time spent on each of the defined academic engagement categories would be most enlightening. Maybe this instrument is not discriminating enough to be used in studies of this type. Maybe the behavior coding procedures need to be delineated in other ways. If the "off-task" category contains a large percentage of the recorded behaviors maybe it needs to be broken down into more definite behaviors. It may also be true that some of the outlined behaviors do not take much more than four seconds to complete. Thus, maybe the four second coding time may be too long. Will students realistically be engaged in a recalling behavior for four seconds? Tables of means, standard deviations, and percentages could shed some light on all these questions.

As mentioned above, the study is basically a strong one establishing or confirming relationships among formal reasoning ability, locus of control, observed students behaviors, and integrated process skill achievement. The relationship between formal reasoning ability and achievement is well established. The Test on Logical Thinking has previously been shown to exhibit significant relationships with achievement in biology and SAT scores (Yeany, Helsett and Barstow, 1980). Similarly, in their instrument validation studies, the researchers reported that in a sample of 353 students in grades six through eight 35% of the variance of integrated science process skill achievement was attributable to scores on the TOLT. However, these findings may not be useful in a realistic sense. This is because, as previously stated, often as few as three or four percent of the sixth and seventh graders score greater than three on the TOLT.

The reported relationship between locus of control and the specific modes of student engagement also could be predicted from general attribution theory. It seems plausible that overt behaviors such as total engagement and attending would be different for internal

and external students. Similarly, the reported relationships between locus of control and achievement are not clear. Some studies find significant relationships and others do not find any relationship. The complex relationship between locus of control and attending behavior may be confounding the issue. That is, attending or total engagement behavior may be more important for some kinds of achievement than others. Thus, if locus of control is highly associated with these stated behaviors, it may not always be positively associated with measured achievement. This study leads to important conjectures in this area.

Lastly, one must be careful not to imply a causal relationship from the correlations among the variables in this study. Although there was a treatment consisting of a sequence of eight science lessons, no attempt was made to see if these lessons in any way affected integrated process skill achievement. The researchers do offer suggestions cautiously, stating that although the results seem to suggest that teachers, in order to increase achievement, could attempt to increase the attending and generalizing behavior of students, several alternatives explanations are equally plausible.

REFERENCES

Campbell, R. L. and L. Martinez-Perez. "Self-concept and Attitudes as Factors in the Achievement of Preservice Teachers." Journal of Research in Science Teaching, 14 (5): 455-459, 1977.

Tobin, K. G. and W. Capie. "The Development and Validation of a Pencil and Paper Test of Logical Thinking." Educational and Psychological Measurement, 41 (2): 413-424, 1981.

Yeany, R., E. H. Helseth, and W. Barstow. "Interactive Instructional Video Tapes, Scholastic Aptitude, Cognitive Development and Locus of Control as Variables Influencing Science Achievement." Paper presented at the annual meeting of The National Association for Research in Science Teaching, Boston, April, 1980.

Lowell, Walter E. "The Development of Hierarchical Classification Skills in Science." Journal of Research in Science Teaching, 17 (5): 425-433, 1980.

Descriptors--*Age; *Classification; *Cognitive Development; Elementary School Science; *Elementary Secondary Education; Learning Theories; *Science Education; Secondary School Science

Expanded abstract and analysis prepared especially for I.S.E by David F. Treagust and Barry J. Fraser, Science and Mathematics Education Centre, Western Australian Institute of Technology.

Purpose

The study explored the development of hierarchical classification and the age at which this appears fully developed in the child. Also, it investigated: (a) the correlation between performances on a battery of nine classification tasks and the performance on the hierarchical classification test; and (b) the effect of IQ and gender on classification ability.

Rationale

The rationale for this study springs from the work of Piaget (1964) and Vygotsky (1962). Piaget believed that mastery of hierarchical ability signals full class reasoning and is a prerequisite to the more complex logical skills found in formal operations. Vygotsky contended that true conceptual thinking cannot proceed until the child is cognizant of the hierarchical nature of concepts. Both Piaget and Vygotsky maintained that the child is capable of this hierarchical ability by 12 years of age. Studies by Kofsky (1966) and Lowell (1978) suggest that hierarchical classification becomes fully developed in late concrete operations. Consequently, the author wished to explore the development of these abilities in detail.

Research Design and Procedures

The study design involved a posttest only with no control group. Data were collected from 120 students randomly selected from five age groups representing primary, junior high and high school grades. Each group of 24 subjects, evenly divided by gender, had a mean age of 6.7, 8.6, 11.6, 13.7, and 16.0 years. Each subject was individually administered:

a. Form 1 of the Quick Test (Ammons and Ammons, 1962) to determine approximate IQ parameters, specifically verbal ability.

b. A set of nine classification tasks, originally developed by Piaget and Inhelder (1964) and Kofsky (1966), to assess seven class operations which constituted a general index of class reasoning ability.

c. A six-level hierarchical classification test (HCT), comprising eight subtests, to measure hierarchical classification ability. This test was designed and has been used by the author in previous studies (Lowell, 1975, 1978).

Findings

Student performance was examined and the results showed that the battery of nine classification tasks was positively correlated with performance on the hierarchical classification test ($r = 0.47$, $p<0.001$). The data show that:

- HCT scores were reasonably stable across similar age groups.
- HCT scores and performance on the class operations measure involving hierarchical classification correlated positively and highly significantly ($r = 0.47$, $p<0.001$).
- The hierarchical classification class operation was the most difficult for children to master.
- Subjects capable of performance on either of the two most difficult tasks from the battery had little difficulty on the remaining tasks.

- Regrouping of subjects according to scores on the hierarchical classification class operation indicates that mastery of this task relates substantially to performance on the HCT.
- There was a significant moderate to strong correlation of mental age with all measured variables.
- Both age and class reasoning ability contributed significantly ($p<0.001$) to hierarchical classification ability.

Interpretations

The findings of the study suggest that:

- Performance on the HCT stabilizes through mid-adolescence.
- The strong positive correlation between the measure of verbal ability and tasks passed and the total number of questions answered on the classification hierarchy suggests that classification is a function of verbal ability.
- Hierarchical classification ability begins to be developed in children at around 11-12 years, but does not become fully developed until mid-adolescence or beyond. The hierarchical classification class operation task was mastered by only 42% of the 16 year olds who were also only performing on the fourth of the six levels of the HCT.
- Of the 18% of the total sample considered to be in the late concrete operations stage, 68% experienced little difficulty on the HCT.
- Piaget's contention is supported, that is, hierarchical classification is not properly understood until late concrete operations.
- IQ and gender have no significant effect on hierarchical classification ability.
- Verbal ability as a measure of IQ is related to all measured variables.

ABSTRACTOR'S ANALYSIS

This clearly written and well organized report describes a carefully designed and executed study of an important area, namely, the development

of hierarchical classification and the age at which it appears. In particular, the study draws upon Piaget's and Vygotsky's earlier work and supports Kofsky's and Lowell's earlier findings.

Although the research design involves a posttest only with no control group, it should be emphasized that this design is typical of Piagetian studies of cognitive performance. Some information supporting the validity of the measures used in the study is provided by the presence of a relatively large positive correlation between the HCT and scores on the classification tasks. Overall the reader would be justified in placing confidence in the validity of the results of this research.

Apart from some confusion caused by the use of abbreviations (namely, HCT and VC) to refer to two distinct but somewhat similar measures, the article is to be generally commended for its careful and complete reporting. An exception to this is the author's lack of attention to referencing details. For example, several authors cited in the text were not in the reference list and some authors in the text had different dates of publication in the reference list.

Although a brief concluding section includes some comments on the study's implications for the science teacher, there is much scope for researchers following the lines of work represented in this paper to make a much more substantial contribution to the translation of research findings into genuinely useful information and suggestions for practitioners.

REFERENCES

Ammons, R. B. and C. H. Ammons. The Quick-Test (QT): Provisional Manual. Missoula: Montana Psychological Test Specialists, 1962.

Lowell, W. "An Empirical Study of a Model of Abstract Learning." Science Education, 61: 229-242, 1977.

Lowell, W. "A Study of Hierarchical Classification in Concrete and Abstract Thought." Journal of Research in Science Teaching, 16: 225-262, 1979.

Piaget, J. and B. Inhelder. The Early Growth of Logic in the Child. New York: Norton, 1964.

Vygotsky, L. Thought and Language. Cambridge, MA: Massachusetts Institute of Technology Press, 1962.

Haladyna, T., R. Olsen, and J. Shaughnessy. "Relations of Student, Teacher, and Learning Environment Variables to Attitudes Toward Science." Science Education, 66 (5): 671-687, 1982.

Descriptors--Attitude Measures; *Classroom Environment; *Elementary School Science; Elementary Secondary Education; Grade 4; Grade 7; Grade 9; Predictor Variables; Science Education; *Secondary School Science; Sex Differences; *Student Attitudes

Expanded abstract and analysis prepared especially for I.S.E. by Marcia Lakes Matyas, Purdue University.

Purpose

The authors examined the "possible determinants of attitudes toward the subject matter of science" (p. 671) by designing a study based upon a theoretical model. This model includes three broad factors which research suggests can influence attitudes: students, teachers, and learning environment. The ultimate objective of the study was "... to better understand the relation of each variable and trait to attitude as well as indicat(e) the extent to which modification of endogenous variables might improve attitude" (p. 674). Specifically, the authors propose three research questions: How much of the variance of a measure of students' attitudes toward the subject of science can be accounted for by exogenous and endogenous student variables and endogenous teacher and learning environment variables?; Are there consistent patterns across grade levels?; and Do gender differences exist in these relationships?

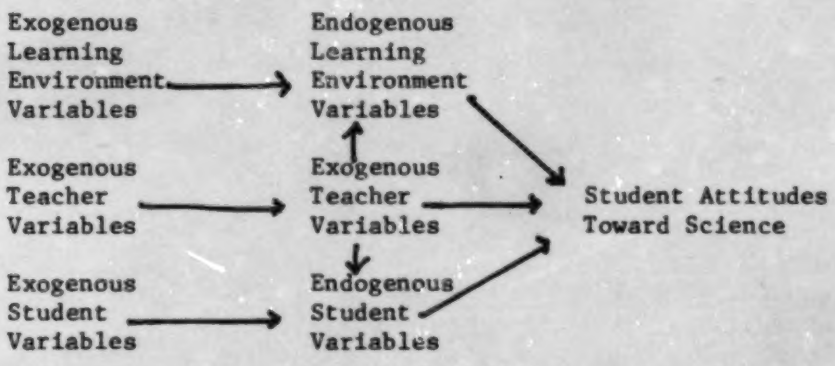
Rationale

A review of literature revealed that several problems exist in much of the research performed on correlates of attitudes toward science. One of the most common problems is the lack of theoretical context for

the studies. In order to avoid this problem, the study proposes the following sequence of action: 1) Accurately describe the dimensions of a model for variables affecting science attitudes; 2) Validly measure the traits and variables included in the model; and 3) Estimate the strength of the relationships of these traits/variables to attitudes toward science.

The model examines two dimensions of predictor variables - content and focus. The variable content can be student, teacher, or learning environment. The variable focus can be either exogenous or endogenous to the organization/institution under study. Table 1 describes the relationship of variables in the model. This model proposes, therefore, that exogenous variables can affect endogenous variables which directly affect student attitude toward science. In addition, endogenous teacher variables can affect both the endogenous learning environment and student variables.

Table 1



Note. Arrows indicate possible effects of one type of variable on another type.

Research Design and Procedures

This study involves non-experimental, survey research techniques and results are, therefore, correlational. A large sample (N=1965) was chosen by stratified random sampling of fourth, seventh, and ninth

grade students in Oregon school districts. The sampling procedure controlled for size of the district and school at each grade level and region of the state. Students were approximately equally distributed among the three grade levels and between males and females.

Data were collected by administration of a single written survey, requiring 30 minutes to complete. All surveys were administered by a trained researcher during regular class time; the teacher was not present during the survey and all responses were anonymous. Partial responses were eliminated from the data analyses.

The survey instrument assessed attitudes toward science, demographic variables, and exogenous student variables as well as endogenous student, teacher, and learning environment variables (Table 2). Survey items were drawn primarily from other research instruments. Construct validity for the survey instrument was reported elsewhere; estimates of content validity were not indicated. The attitudes toward science subscale consisted of four questions; scoring methods and reliability estimates were not given for this subscale. Reliability estimates for each of the other subscales, listed in Table 2, ranged from .05 to .94. However, several variables were assessed by single items on the survey and, therefore, no reliability estimates were determined for these variables.

Table 2

Exogenous Student Variables: Family mobility; amount of TV viewing; School attendance; Parental involvement; Parental concern about school

Endogenous Student Variables: Academic self-concept; Fatalism; Concern for grades; Importance of science; Amount of homework done

Endogenous Teacher Variables: Overall teacher quality; Teacher enthusiasm; Respect for teacher; Teacher support for the individual; Praise and reinforcement; Commitment to learning; Fairness to students

Endogenous Learning Environment Variables: Satisfaction; Enjoyment of classmates; Friction in class; Difficulty of subject; Cohesiveness; Competition; School environment; Homework assigned; Formality; Speed; Class environment; Attentiveness; Goal direction; Favoritism; Cliqueness; Disorganization; Apathy; Diversity; Materials Usage

Data were analyzed through use of product-moment correlations and stepwise multiple regression analyses. Both statistical ($p < 0.01$) and practical ($r = .20$) criteria were utilized to determine the significance of these correlations. Multiple regression analyses were conducted separately for males and females within each grade level; significance level for inclusion of variables in the equations was $p=0.05$.

Findings

Correlational analyses

Exogenous student variables. Non-significant correlations occurred in all but one case: ninth grade boys' parental involvement was significantly related to attitude.

Endogenous student variables. Self-confidence was positively related to attitude ($r = .23$ to $.41$), with the highest correlations occurring among ninth grade students. Fatalism was negatively related to attitude ($r = -.26$ to $-.53$) and students' perception of the importance of science was positively related to science attitudes ($r = .31$ to $.52$), especially among seventh and ninth grade students. Neither student concern for grades nor amount of homework completed was highly related to science attitudes.

Endogenous teacher variables. For each of these variables (except teacher enthusiasm), lower correlations with science attitudes were found among students in the fourth grade ($r = .12$ to $.22$) than among those in seventh ($r = .31$ to $.55$) and ninth ($r = .36$ to $.55$) grades. Also, for each endogenous teacher variable (except fairness), correlations for female students were "observably larger" than those for male students at the same grade level.

Endogenous learning environment variables. Enjoyment of classmates was moderately related to attitude for students at all three grade levels ($r = .24$ to $.46$). Satisfaction with the class was related to attitude ($r = .31$ to $.43$) for all groups except fourth grade girls.

and seventh grade boys. Class environment and attentiveness were positively related to attitude among both seventh and ninth grade students (r = .30 to .48) while disorganization was negatively related for both groups (r = .27 to -.35). Materials use, apathy, goal direction, favoritism, and diversity were not clearly related to attitude.

Regression analyses. Regression equations for each of the four content/focus variable sets (listed above) were conducted using science attitude scores as the criterion variable. A "best" model analysis using variables from all four content/focus groups in the regression equation was then performed.

Fourth grade students. None of the regression equations on the four sets of variables accounted for a sizeable proportion of the variance in attitude scores (8.7% to 20.1%). In the best model analysis, satisfaction with the class was the most important variable for boys, accounting for 11.7% of the variance in attitude scores. For girls, fatalism entered the equation first, accounting for 14.6% of the variance. Total accounted variance for boys and girls with the best model was 24.9% and 24.2%, respectively.

Seventh grade students. Both fatalism and the importance of science were significant variables in equations for both males and females, accounting for approximately 30% of the variance in science attitude scores. Teacher quality was a significant predictor for both males (22.7% of variance) and females (30.1% of variance). Among the learning environment variables, classroom environment accounted for the greatest amount of variance among both males (15.8%) and females (21.0%). The best model accounted for approximately 40% of the variance in attitude scores for both boys and girls; importance of science was the best predictor for boys while teacher quality was the best predictor for girls.

Ninth grade students. Among the student variables, fatalism and the importance of science accounted for the most variance in attitude scores among males and females (27.6% and 39.4%, respectively).

Teacher quality was the best predictor for both boys (22.5%) and girls (30.7%) among the teacher variables. The best model equation for ninth grade students was similar to that for seventh grade students.

Interpretations

Conclusions

1) There was "...no evidence to support any meaningful relationships between exogenous student variables and science attitudes" (p. 658).

2) "Student attitudes toward science are powerfully related to both students' sense of the importance of science and their level of fatalism" (p. 685).

3) There is a "consistently high relationship" of endogenous teacher quality variables to science attitudes.

4) Endogenous learning environment variables are "highly related" to science attitudes. However, few clear patterns are evident; only class environment and organization consistently were related to science attitudes.

Inferences and implications. The authors suggest that the model should be further tested by causal analyses (path analysis or LISREL) and longitudinal, experimental studies. They further state that the proposed model "...implies that efforts to strengthen teacher quality and the learning environment will have a positive effect on student outcomes such as attitudes" (p. 686).

ABSTRACTOR'S ANALYSIS

Theoretical context. This study approaches the investigation of relationships of attitudes toward science to other variables from a theoretical perspective. The authors have proposed a model which classifies variables according to their content and focus, and have investigated several variables in four of the six possible content/focus

areas. As noted by the authors, science attitudes research frequently fails to design studies based upon a theoretical model, therefore, the proposal of a theoretical model for factors affecting science attitudes among elementary and junior high school students is a significant contribution. While the proposed model is a fairly simple one, this study provides a theoretical basis and initial data which can serve as the groundwork for further testing, modification, and expansion of the model.

One possible expansion of this model is the examination of additional relationships between the groups of variables. For example, while the model proposes that endogenous teacher variables may affect learning environment variables, it may also be possible that certain endogenous learning environment variables may affect teacher variables. Another possible expansion which could be easily tested is the direct (as opposed to the suggested indirect) effect of exogenous student variables on attitudes. Although a few exogenous student variables were tested, additional variables (for example, extracurricular science experiences, parental perceptions of the importance of science, parental perception of the child as a science student, exposure to science role models, etc.) might prove to be more closely related to attitudes toward science than were the exogenous variables examined here.

Validity and Research Design. While the procedures used for sample selection and the administration of the survey instrument were appropriate, several questions can be raised about the survey instrument itself and the data describing it.

1) Several (seven) variables were assessed by single items on the survey. Most notably, the measurement of the student's perception of the importance of science was assessed by a single question. Results indicate that this variable was one of the most important predictors of positive science attitudes for both males and females at each grade level. What can be stated about the reliability and validity of measurements using single items?

2) The reliability of several of the multiple-item subscales is questionable. Nearly 30% of the reliability estimates for the subscales of the survey were less than 0.50 and some were as low as 0.05. What can be done to improve the reliability of these measurements?

3) Was a concise and clear description of efforts to establish the scale's validity possible? The number of items comprising each subscale of the survey was not indicated nor was any mention made of content validity estimates for the subscales. Furthermore, the reference describing the construct validation of the instrument was a paper presented at a national meeting; this type of document is not easily accessible nor is it always subject to written peer-reviews. A brief description of the content validation procedure and findings would be helpful.

4) How accurate and valid was the measurement of the criterion variable: attitudes toward science? The subscale measuring attitude toward science consisted of only four questions and no estimates of reliability or validity were given.

5) Why were the particular demographic variables (family background, mobility, and socioeconomic status; TV viewing; and school attendance) chosen for comparison of groups (grade level and gener)? Could quantitative (chi-square) comparisons of these groups have been performed? How was socioeconomic status measured? Could family background, socioeconomic status, and attendance data be compared to state or national data in order to make inferences about generalizability of the study?

The methods chosen for data analysis were appropriate for the type of data collected and the purposes of the study. In particular, the use of both statistical and practical criteria for "significance" of product-moment correlations is commendable. It would also have been helpful to establish practical criteria for the percent accounted variance in the multiple regression analyses and for gender differences in product-moment correlations.

Problems in attitude research. The authors specifically cite several problems in research on attitudes toward science. It is appropriate, therefore, to examine whether this study did or did not avoid these problems.

1) "Most studies have lacked a theoretical context" (p.671). As stated earlier, this study proposes a theoretical model for the relationship between endogenous and exogenous teacher, student, and learning environment variables and student attitudes toward the subject of science. Although this study did not specifically test the model, it does provide data which might be suitable for further analyses and model tests.

2) "Definitions of science attitudes vary greatly and are generally confounded" (p. 671). "Attitude toward science" is neither theoretically nor operationally defined in this study. A theoretical definition should be included in the description of the model and an operational definition, including estimates of validity and reliability, should be included in the description of the instrument.

3) "There is a lack of integrative research findings in the literature" (p. 671). The authors propose a model which includes variables from three major content areas: student variables, teacher variables, and learning environment variables. It should be noted, however, that the authors did not measure or give examples of exogenous teacher and learning environment variables which might be related to attitudes toward science. How these factors actually fit into the proposed model is, therefore, somewhat unclear.

4) "The selection of variables induced in studies seems haphazard and asystematic" (p. 671). Although the selection of the three general content areas was well-justified by citation of previous research findings, the rationale for choosing the specific variables included in the study was not explicitly stated. Previous research findings which suggest a relationship among these specific variables also might be cited. This is particularly true for exogenous and endogenous student variables and endogenous teacher variables since a wide variety of variables in these content/focus areas have been related to science attitudes.

5) "Development and validation of suitable instrumentation for attitude measurement is greatly needed" (p. 671). As discussed earlier, several questions exist concerning the survey instrument.

However, the development of this instrument appears to be a positive step toward the ultimate goal of developing a "suitable" science attitudes scale. If both validity and reliability estimates can be firmly established, the present instrument may provide an excellent method for assessing student attitudes toward science as a subject.

6) "Improvement is needed in both the design of studies and the analysis of data. The univariate and non-statistical approaches commonly in use provide little in the way of usable information for science educators" (p. 671). The use of appropriate sampling procedures and the application of practical criteria to the analysis of correlational data indicate adherence to good research design principles. Unfortunately, the analysis of gender differences was primarily qualitative and non-statistical. If conclusions are to be drawn from an analysis, descriptive terms such as "not nearly so strong" (p. 678), "approximately the same strength" (p. 680), "observably larger" (p. 681), and "much lower" (p. 681) must be avoided. When quantitative statistical analyses are not feasible, then practical criteria could be established for comparison of product-moment correlations.

Suggestions for future research. As suggested by the authors, causal analysis (path analysis or LISREL) would be appropriate subsequent steps in the testing of their model. Due to gender and grade level differences found in this study, separate analyses for males and females within each grade level would be warranted. In addition, the possiblity of a longitudinal study to determine attitude changes across grade levels should be explored. The authors of this study suggest that the negative association between attitude toward science and fatalism "apparently grows stronger with passing years" (p. 680); this type of conclusion must be drawn with great caution without supporting data from longitudinal studies. Finally, as suggested earlier, this theoretical model could be expanded to include additional relationships between variable groups and future research could include exogenous teacher and learning environment variables.

Padilla, Michael J. and LLoyd Ollila. "Effects of Small-Group Teaching on Acquisition and Transfer of Nonvisual Seriation Abilities." Science Education, 64 (3): 357-366, 1980.

Descriptors--Elementary Education; *Elementary School Science; Grade 1; *Learning Theories; Science Education; Science Instruction; *Serial Ordering; *Small Group Instruction

Expanded abstract and analysis prepared especially for I.S.E. by Gerald G. Neufeld, Brandon University, Canada.

Purpose

The purpose of this study was to determine the relative effectiveness of two small-group instructional strategies on first-grade children's acquisition, retention, and transfer of seriation abilities.

Rationale

Research on the seriation strategies used by young children, (Smith and Padilla, 1977 and Baylor and Gasoon, 1974) indicates that there are two distinct strategies children use in successfully solving seriation tasks. Subsequent research (Padilla and Smith, 1979) has indicated that first-grade students can learn to use these strategies when individually trained and that the strategy involving finding and eliminating the extreme object was particularly helpful for stage I children. Other researchers, (Schafer and Byers, 1975 and Padilla and Smith, 1979) found that individual training using corrective feedback, attention to task stimuli, and other instructional methodologies produced substantial and durable seriation behavior change. The "elimination" and "corrective feedback" instructional strategies formed the basis for the treatments used in this study.

The previous research on teaching children to seriate has used one-to-one instruction. As this form of instruction is not practical in a normal elementary school classroom, this study attempted to replicate the previous findings using small group instruction.

Research Design and Procedures

One-hundred and twenty first-grade students from two average Victoria, British Columbia, public schools were individually tested using the stick task (Inhelder and Piaget, 1964). There were 25 stage-I, 57 stage-II, and 38 stage-III children. Twenty-four children from each stage were randomly selected for inclusion in the study. Within each stage, six groups of four students were randomly selected. Two groups from each stage were randomly assigned to the three treatment conditions: Extreme Value Selection (EVS), Control (CON), and Placebo (PLAC). The groups were randomly assigned to one of three trainers.

Seriation training and testing was done using sets of materials that varied only in nonvisual variables. Eight cups ranging in mass from 10 to 2100 grams were used for ordering by mass. Eight cylinders lined with materials ranging from smooth acetate to 40 grit sandpaper were used for ordering by texture. To test for transfer on the post-test and retention test a set of eight pipes with handles that took from 0.8 to 11.25 lbs. to pull was used for ordering by force.

The experimental treatment for each of the 18 groups of four students consisted of four training sessions. Each session lasted approximately 15 minutes and was separated by one or two school days. The sessions alternately employed the weight and texture materials.

Each session for an Extreme Value Selection (EVS) group began with an introduction to the materials and the task. Each child was asked to observe a correctly ordered set of four objects as the task was described. The EVS selection was then modeled and orally described in steps. After each step, the children were asked to perform the same action on their materials. Corrective feedback was given. After successfully ordering three objects, the students were encouraged to finish the task. For the remainder of the session the children practiced the task and were given feedback. Individual assistance was given to those needing it.

Each session for a Control (CON) group also began with an introduction to the materials and the task. Each child was asked to

observe a correctly ordered set of four objects as the task was described. The children were then told to "...put them in a row from heaviest to lightest (or roughest to smoothest)." When they were finished they were given feedback as to correctness and told to continue practicing.

The Placebo (PLAC) children were given no introduction to the task or materials. They were merely encouraged to observe and explore their materials in creative ways, some of which might involve observing the variables of interest and actually ordering on those variables.

Within two to three school days of the last training session, all children were individually posttested. The posttest consisted of one chance to order the weight, texture, and force materials as well as the stick test (Inhelder and Piaget, 1964) that had been used as a pretest. Approximately three months later these same tests were repeated as a retention test.

A task accuracy score (TS) was calculated for each task on the pre-, post-, and retention tests. The score was the Kendall's tau correlation between the child's ordering and the true order.

The three trainers spent eight workshop hours practicing the training procedures to a mastery level. In addition, they followed a protocol for each of the instructional treatments. Testers received about four hours of training. All quickly reached and maintained 95% concurrence with each other and a master trainer. No trainer tested students from treatment groups that he/she taught. Face validity for the measured is claimed.

A 3x3 factorial ANOVA design was used to analyze the results. The three treatments (EVS, CON, and PLAC) were fully crossed with the three stages (I, II, and III). An alpha level of 0.05 was used for significance testing. However, probabilities up to 0.10 were reviewed since defining and inferring relationships was the major intent of the study.

Findings

The confidence intervals (97.5%) for the means of the TS scores overlapped for the comparable groups in the present study and the earlier study (Padilla and Smith, 1979) for both the post- and retention tests. The authors indicate that this demonstrates that small-group instructional methods, as used in the present study, can be as effective as individual training.

When the posttest data from the present study were analyzed using a 2-way ANOVA, a significant interaction effect ($p<0.02$) as well as a main effect ($p<0.001$) was found. Further analysis indicated that both EVS and CON treatments were more effective than PLAC and that EVS was more effective than either CON or PLAC with Stage-I children.

An ANOVA analysis of the retention test data indicated a significant effect due to treatment ($p<0.01$). Further analysis indicated that training was retained by all groups except the EVS Stage-I children, whose scores dropped considerably and showed a substantial increase in variability.

An ANOVA analysis of the posttest transfer task scores indicated a significant difference due to treatment ($p<0.01$). Further analysis indicated that EVS children scored higher than PLAC children ($p<0.10$). An ANOVA analysis of the retention test transfer task scores indicated a significant difference due to treatment ($p<0.001$). Further analysis indicated that CON children scored higher than PLAC children.

The stick test (Inhelder and Piaget, 1964) was used as part of the pre-, post-, and retention tests. The results of the stick test did not seem to favor any particular treatment. Stage-III children maintained their ordering and insertion abilities. Stage-II children performed consistently on the ordering task and showed increasing competence on the insertion task. Stage-I children showed increasing competence on both tasks although all treatments showed a high variability.

Interpretations

On the basis of their findings the authors concluded that:

a) small group teaching methods as used in this study can be as successful as individual training.

b) although the EVS treatment was more effective in producing short-term accuracy in those children who do not have the necessary logical structures (Stage-I children), this treatment appears to create only rote learning which is lost over time, and

c) the CON treatment, which provided only "correctness" feedback and thus forced the children to reorganize their own structures to correctly solve the task, produced long-term learning. This lends strong support to Piaget's equilibration model of learning and emphasizes the value of guided-discovery science teaching strategies.

ABSTRACTOR'S ANALYSIS

This study by Padilla and Ollila is a logical extension of the existing body of research on the learning of logical operations such as seriation. A particularly valuable aspect of the study is the authors' use of treatments that can be readily implemented in normal classroom situations.

In general, the study was quite well designed. The sample size (72 children) was too small to detect relatively small effects. Few educational research studies are able to use random assignment of subjects to treatments. Padilla and Ollila are to be commended for taking the time and effort necessary to arrange for random assignment of subjects to treatment groups. This made their research design much more rigorous and defensible.

Before the children were randomly assigned to groups, they were pre-tested and classified as Stage-I, II, or III seriators. Since the

objective of the instructional treatments was to teach the children how to seriate, one wonders why the Stage-III children were included in the study. These children were the ones who were already able to seriate and had obtained a perfect score on the pre-test. Further instruction in seriation would seem to be pointless for these children. The inclusion of these students in the study merely guaranteed that the scores would not be normally distributed and, as a result, weakened the statistical analysis.

Although the basic design of the study was quite good, some questions can be raised regarding the possible effect of other instruction in seriation the children may have received in school. Many first-grade language arts programs include instruction and activities in ordering. For example, as a pre-reading activity a child might be given a series of pictures showing a seed, a seedling with one leaf, a seedling with two leaves, a plant with four or five leaves, a blooming plant, and a plant with seeds and asked to place them in order. Since most first grade teachers divide their classes into groups based on reading readiness, these seriation activities may have occurred between the pre- and retention tests for some students and not for others. If this happened, the results of the study would be contaminated. The research report does not indicate whether or not this may have occurred.

The authors provided a good description of the dependent variables used and how the task accuracy scores were calculated (a Kendall tau correlation coefficient between the child's ordering and the true order). Unfortunately, no information or references to the statistical research literature are given to justify their subsequent statistical treatment of these task accuracy scores. It would appear that these scores represent measures on an ordinal scale. As a result, there seems to be little statistical justification for their analysis of these scores. You must have parametric measurements based on interval or ratio scales before means, standard deviations, and analyses of variance have any real meaning. This inappropriate use of statistical techniques may be what caused the ANOVA analysis of the posttest

transfer task results to indicate a significant difference between treatments ($p<0.01$) while subsequent comparisons of means failed to detect any significant differences at the 0.05 level. It would seem that nonparametric methods of analysis would have been more appropriate for this study.

The research report included a good description of: the subjects and how they were selected, the experimental treatments used, the dependent measures used, and the methods used for analyzing the results. It would have been improved if the authors had:

a) provided a more substantial rationale for the exclusive use of instructional materials that varied in nonvisual characteristics,
b) used a name other than Control for the experimental group which received only feedback on correctness (the word "control" has very specific implications),
c) briefly described the statistical procedure used for the post hoc analysis of differences between means, and
d) been more careful in drawing conclusions (you can never "accept" an hypothesis, you can only "fail to accept" the alternative).

However, the research report, as a whole, was very well and carefully written. As a result, it should be relatively easy for an interested researcher to replicate the study.

Although the results of the study are interesting and suggestive, they hardly justify the authors' statement that "... the results with the CON treatment lend strong support to Piaget's equilibration model." Since a child spends approximately 12,000 hours attending public school, educational researchers are deluding themselves if they expect to produce significant changes in a child's intellectual growth or provide strong evidence for or against a broad theory of learning by conducting even a well-designed experiment in which the experimental treatment lasts for one hour. Although small scale educational experimentation is useful, we will not make major strides in understanding or changing the intellectual development of children unless we begin to conduct long term studies that involve modifications of significant portions of a child's education.

REFERENCES

Baylor, G. O. and J. Gasoon. "An Information Processing Theory of Aspects of the Development of Weight Seriation in Children." Cognitive Psychology, 6: 1-40, 1974.

Inhelder, B. and J. Piaget. The Early Growth of Logic in the Child. London: Routledge and Kegan Paul, 1964.

Padilla, M. and E. Smith. "Experimental Results of Teaching First Grade Children Strategies for Nonvisual Seriation." Journal of Research in Science Teaching, 16 (4): 339-345, 1979.

Schafer, L. and J. Byers. "The Effectiveness of Cue Fading in Teaching Kindergarten Children to Serial Order." Journal of Research in Science Teaching, 12: 281-292, 1975.

Smith, E. and M. Padilla. "Strategies Used by First-Grade Children in Ordering Varying Numbers of Objects by Length and Weight." Journal of Research in Science Teaching, 14 (5): 461-466, 1977.

Mayer, Victor and M. James Kozlow. "An Evaluation of a Time-Series Single-Subject Design Used in an Intensive Study of Concept Understanding." Journal of Research in Science Teaching, 17 (5): 455-461, 1980.

Descriptors--*Concept Formation; Earth Science; *Measurement Techniques; Science Education; *Science Tests; *Scientific Concepts; Secondary Education; *Secondary School Science

Expanded abstract and analysis prepared especially for I.S.E. by Frances Lawrenz, Arizona State University.

Purpose

The purpose of this research was to further the development of time series designs for use in science education, in concept learning in particular. The authors investigated the degree of classroom interference associated with administering a daily one-item or three-item test and the validity of the two testing methods.

Rationale

The rationale for the study was that a time series design had previously been shown to provide valid information on science attitude development and that a logical next step was to investigate the usefulness of time series in examining concept development.

Research Design and Procedure

The authors selected two eighth grade earth science classes as the subjects. The students were in a junior high school in a central Ohio suburban community with a broad socioeconomic range. The two classes were taught by the same teacher. His teaching style was informal with almost all instruction of a student centered, small group, laboratory or activity oriented nature.

The intervention was a 10 day unit on crustal evolution. This unit was developed around three major objectives and an item pool measuring understanding of all of the concepts involved was generated. Both the instructional unit and the items were field tested and modified accordingly. After the item analysis and revisions there was an item pool of 54 multiple choice items.

The students were tested daily for 26 days. The first eight days constituted the baseline stage. The next 10 days were the intervention, the unit on crustal evolution, and the last 8 days were follow-up. One class, A, responded daily to a one item instrument with the item randomly drawn from the item pool. No student received the same item more than once. The second class, B, responded daily to a three item instrument, with one item randomly selected from each of the three unit objectives. Every student in class B completed the same test on a given day and different tests were given on successive days. Both classes completed 32-item posttests at the end of the intervention. There were 10 items common to both forms of the posttest. The KR 20 reliabilities for these tests were: class A's, .77 and class B's, .71. In addition to the daily and posttesting, both classes were observed every day during the intervention to verify the similarity of the two classes' experiences and to record the amount of time taken for the administration of the tests.

The research design was two single subject time series with intervention. The data from all the students were reduced to a single score for each class for each day. For class A, which received the one item test, the number of students who responded correctly to their item was divided by the total number of students present. For class B, which received the three item tests, the student scores were first converted to a proportion by dividing the number correct by three and then the class mean was computed. Data from students who missed three or more days of the ten day intervention were omitted from the analysis. The resulting data were graphed and the patterns for the baseline, intervention, and follow-up stages were compared. Correlations were calculated between day and class score for each stage. The effect of the difficulty levels of the items was also investigated using item difficulty indexes from the two posttests.

Findings

The log of classroom observations indicated that the one-item testing procedure took an average of less than two minutes to administer and the three item instrument only slightly longer. Since the tests were administered daily, they became part of the classroom routine and resulted in minimum disruption.

The authors expected the data to follow a traditional learning curve where the amount of information learned increases over the period of time instruction takes place and then drifts down due to forgetting after instruction stops. Examination of the graphed data showed an increase in scores for both class A and class B during the intervention and then a decrease during the follow-up stage. The correlations showed no significant relationships between day and class score for the baseline stage for either class. The correlation for class B was positive at the .1 level, but both classes seem to have exhibited a spuriously high score for the first day of the baseline stage. The scores for both classes showed a significant ($p \leq .05$) positive correlation with time during the intervention. During the follow-up stage the scores for both classes showed nonsignificant negative correlations with day.

The scores for the first two days of the follow-up period for both classes (those days right after the intervention) were quite high suggesting a "momentum effect." Therefore the correlations were recalculated including these two days with the intervention and having only 6 days as the follow-up stage. With this grouping the correlations of scores with day for the intervention were both positive and higher than for the 10 day intervention period. The correlations of score with day for the follow-up period were still nonsignificant although they were now positive rather than negative.

The effect of difficulty of the items was examined since the classes did not perform at the same level (class A, $\overline{X} = 6.11$; class B, $\overline{X} = 4.67$) on the 10 common items in the posttests. Class B performance was chosen as the standard of item difficulty and the relative difference between

class A and B was used to adjust the difficulty levels of the items that class B did not take. Using these calculated difficulty indexes, the range of difficulties for class A was .43 to .56 but for class B was .30 to .66. Because only 18 of the three item tests used in class B were developed, the students took the same three item tests during the baseline and follow-up stages. The authors suggested that perhaps the tests used at the end of each stage were less difficult than the tests used at the beginning of each stage and therefore the students got higher than expected scores at the ends of the two stages.

Interpretations

The authors suggested that the validity of the one item technique for obtaining data on knowledge acquisition was demonstrated because the pattern of the data from class A followed that of a standard learning curve. Data from class B, however, did not follow the standard learning curve pattern and the authors suggested that this might be due to the fluctuation in difficulty level from day to day with the three item test for the entire class approach to gathering the data.

The authors discussed the possibility of an effect due to familiarization with the items. This was discounted on the grounds that no one in class A received the same item twice and that, although those in class B did receive some items twice, their greater decline in performance during the follow-up stage indicated that they did not learn the items. Further, it was thought unlikely that the students from either class would have discussed the items among themselves since the items were not used for grading purposes and no discussion of items was noticed by the observer.

In summary, the authors stated that the study demonstrated that a data-collecting procedure can be developed for measuring concept understanding in an intensive single subject time series design. The results indicated that, although neither method caused much classroom disruption, the one item per student collection procedure was more valid than the three items per class procedure because of the fluctuation in difficulty levels.

ABSTRACTOR'S ANALYSIS

This report is well written and succinct. The authors were very thorough in their methodology. Their subjects, instruments, and procedures are well described although a few sample items or a statement of the three major crustal evolution objectives would have been a nice addition. The Tables and Figures are clear, well labeled and provide all of the necessary information.

The authors link their research to only one other study, Mayer and Lewis (1979). This is not surprising since little research has been done using time series designs in science education. Both of these studies represent a new and potentially fruitful method for examining science education. Some of the potential advantages are pointed out in the Mayer and Lewis study, e.g., smaller samples and many data points. The present study, however, does not elaborate on these advantages nor do the authors make explicit connections between their findings and science education research. A short discussion on the future use of time series designs would have been beneficial. Certainly the authors know more about the potential advantages and disadvantages of time series designs than many others in the science education field and their opinions would have helped to direct future studies. This is especially true since the authors themselves consider their study to be an exploration into the usefulness of time series design.

The authors were comparing two data collection procedures; one that allowed for one item per student and one that allowed for three items per class. The authors, do not, however, discuss why these two methods were chosen, or their advantages, or when one or the other procedure might be more appropriate. The authors concluded that the three item method was less valid because of fluctuations in the daily difficulty level since the whole class took the same three items. There was little comparison of the effect of the length of the tests. Apparently tests of either one item or of three items in length (or of the items as in the Mayer and Lewis study) cause a minimum of disruption and provide valid data. The problem with validity arises

when the entire class takes the same items on a given day because then the difficulty level may vary. Naturally this problem could be alleviated by developing tests for each day of the same difficulty level, if it were determined that it was beneficial for the entire class to take the same items. Since no background on the appropriate use of either of the two testing procedures was given, it is difficult to judge when a researcher might want to have the entire class take the same items.

The idea of using time series designs for science education research is exciting. It can provide a new perspective to research and permit a finer grained analysis of what goes on in the classroom. This study increases the likelihood of using time series by demonstrating that it can be used to assess concept knowledge acquisition as well as attitude development and by presenting two methods of data collection. A researcher wishing to employ a time series design now has a model to follow.

There are some limitations with using the time series model. Both this study and the Mayer and Lewis study were testing the feasibility of time series design. Neither study used the design to investigate issues in science education. Both studies pronounced the time series approach as valid since the results were consistent with other research or theories in the area. In this study, for example, the method was deemed valid because a reasonable facsimile of a traditional learning curve seemed to approximate the data obtained from class A and extenuating circumstances were cited for the seeming discrepancies in the class B data. Similar recourse to traditionally held opinions about what attitudes should be supported the validity of the time series method in the Mayer and Lewis study. It does seem that the approach is valid for these purposes. More conclusive evidence, however, would be provided if two experiments were performed simultaneously using comparable groups and the same test items and varying only the design, i.e., time series versus pre, post and delayed posttesting. In this way the advantages of each technique could be delineated and the validity of the time series approach made more explicit.

Another limitation of the time series approach is the type of analyses that can be completed. Traditionally, time series are examined for trends and regularities and modeled by some function. In an interrupted time series design the original function modeled on the baseline data is compared to the new function modeled on the series obtained after the intervention. By comparing these two functions an estimate of the effect, if any, of the intervention can be obtained. These methods are well defined and reasonably rigorous although intuition does play a larger role than in other research methods. The analyses applied in this study consisted of looking at the graphed data and calculating correlations of day with class score. Both of these methods are quite subjective. The correlations with their significance levels provide some quantification, but the tendency to make judgements based on even nonsignificant correlations is strong. Even in this study the authors state that the nonsignificant but negative correlation for the follow-up stage for class A supports an example of forgetting while the nonsignificant but negative correlation (albeit lower) for class B does not. On what basis was the decision made if neither correlation was significant? Looking at the graphed data can also be misleading and very easily biased in favor of the expected outcomes.

The possiblitiy of biased interpretations is increased as the number of observations decreases. In this study eight data points made up the base line stage. Eight is a very small number. A more traditional approach to time series would require a minimum of 50 observations to meet the theory assumptions. The authors do not justify their use of 26 data points to examine three essentially different series. This type of justification should be provided in future studies. Further, future research should compare the results obtained from time series with both large and small numbers of observations to accurately determine the effect.

In summary the study was a good second step (the Mayer and Lewis article being the first) toward the incorporation of time series into science education research. It has been shown that time series data can be gathered efficiently for both attitudes and concept knowledge and

that the results obtained from this data gathering are consistent with expectations. In the future, however, a more rigorous testing of the model needs to take place. Results from the time series approach need to be carefully compared with results from other approaches. The critical number of data points necessary for valid data needs to be determined and the data obtained should be subjected to more quantitative analyses.

REFERENCES

Mayer, V. J. and D. K. Lewis. "An Evaluation of the Use of a Time Series Single Subject Design." Journal of Research in Science Teaching, 16: 137-144, 1979.